Eur J Vasc Endovasc Surg **30**, 215–218 (2005)
doi:10.1016/j.ejvs.2005.01.032, available online at http://www.sciencedirect.com on SCIENCE DIRECT®

Setting Standards for the Assessment of Operative Competence

J.D. Beard on behalf of the Education and Training Committee of the Vascular Society of Great Britain and Ireland

Sheffield Vascular Institute, Northern General Hospital, Sheffield, UK

Background. *We have previously shown that the operative competence of trainees can be reliably assessed using structured checklists and video recording. These assessments are useful for training (formative assessment). However, a standard setting exercise is required before they can be used for examinations (summative assessment).*
Methods. *Blinded videos of a saphenofemoral disconnection by an experienced (competent) trainee and an inexperienced trainee (not competent) were scored, using a structured checklist, by 14 consultant vascular surgeons and 14 vascular trainees. The observers were also asked to decide whether the surgeon was competent, borderline or not competent. Thirteen vascular operating room (OR) nurses performed the same exercise. The 'contrasting groups' method was used to compare the cut point between the scores.*
Results. *There was complete separation between the surgeons' scores for the experienced trainee (median 16, range 13–18) and the inexperienced trainee (median 6.5, range 2–12), Chi-square* $p = 0.0001$. *This separation was confirmed by the judgements for competent (14–18), borderline (15–7), and not competent (8–2),* $p = 0.0001$. *Trainees awarded lower scores than the consultants to both videos, although this difference was not significant. The nurses performed almost as well as the surgeons.*
Conclusions. *Surgeons can discriminate between the video recordings of a competent and non-competent trainee. Such recordings could form part of a trainee's portfolio and contribute to subsequent examinations (summative assessment). It seems that nurses can also be used to assess the operative competence of surgical trainees.*

Keywords: Operative skills; Competence; Assessment; Video recording; Standard setting.

Introduction

As well as operative skills, surgeons require knowledge, communication skills, judgement and the capacity for focused and sustained attention.[1] However, good operative skills are the bedrock of successful surgery. It seems surprising, therefore, that most surgical examinations have not assessed these skills. The reasons for this are that it seemed easier to test knowledge and the long duration of conventional surgical training usually ensured operative competence. We now know that it is important to test all the skills required of a surgeon, especially because of the reduction in the hours available for surgical training.[2,3] The European Board of Vascular Surgery has recently introduced tests of operative competence, using simulations including saphenofemoral disconnection (SFD), into the EBSQ-VASC examination.[4]

The operative competence of trainees can also be assessed in the operating room (OR) by direct observation or video recordings.[5,6] We have shown in a previous study that assessment of operative competence by direct observation or video recordings, using structured checklists (OSATS), has excellent criterion validity (i.e. experienced trainees achieve higher scores) and inter-observer reliability.[7] Video recordings combined with such checklists provide valuable feedback to the trainee (formative assessment). Some specialities, e.g. primary care, use such video recordings for their examinations.[8] However, a standard setting exercise, to establish the cut-point between pass and fail, is required before such recordings can be used for examination purposes (summative assessment).

Methods

Videos of a SFD performed by an experienced vascular trainee (106 SFDs previously performed) and an inexperienced trainee (only nine SFDs previously performed) were scored by 14 consultant vascular

Presented at the Annual Meeting of the European Vascular Surgical Society, Innsbruck, September 2004.
E-mail address: jonathan.d.beard@sth.nhs.uk

1078–5884/000215 + 04 $35.00/0

surgeons and 14 vascular trainees. The observers used an 18-point binary checklist, combining task-specific and global items,[6] which had been validated in a previous study.[7] The videos were anonymised with regard to the experience of the trainee. Fast-forwarding was permitted during uneventful sections of the video. The observers were also asked to judge whether the surgeon was competent, borderline or not competent to perform the operation. Thirteen vascular OR nurses performed the same exercise.

The 'contrasting groups' method was used to compare the cut point between the scores.[9] This method is one that is used to set the pass/fail mark for examinations. The distribution of scores for those judged to have passed (equivalent to competent) are compared to those judged to have failed (not competent). The pass/fail mark is set at the point where the two distributions cross (the cut point). An ideal test is one where there is complete separation between the scores, but this is unusual.

Results

There was complete separation between the scores given by the vascular consultants and trainees for the experienced trainee (median 16, range 13–18) and the inexperienced trainee (median 6.5, range 2–12), Chi-square $p<0.0001$ (Fig. 1). This separation was confirmed by their judgements for competent (14–18), borderline (15–7), and not competent (8–2), $p<0.0001$. The majority of surgeons rated the experienced trainee as competent and the inexperienced trainee as not competent. Four surgeons rated the experienced trainee as borderline and 12 rated the inexperienced trainee as borderline but, none rated the experienced trainee as not competent or the inexperienced trainee as competent (Fig. 2). The trainees scored both videos more harshly than the consultants (median for experienced trainee 15.5 (range 13–18) compared to 17 (range 15–18) and median for inexperienced trainee 4.5 (range 2–12) compared to 8 (range 5–12), respectively. However, this difference did not reach significance (Fig. 3).

The vascular OR nurses performed almost as well as the surgeons (Fig. 3). There was no overlap between their scores for the experienced and inexperienced trainee (median 18, range 14–16 compared to median 7, range 2–14, respectively) and the difference remained highly significant ($p<0.0001$). Their judgements regarding competence were also similar to the surgeons.

Discussion

Standard setting exercises are common practice in knowledge-based examinations,[9] but this is one of the first studies to apply the same methodology to a skill-based assessment. The results demonstrate that blinded surgical assessors can discriminate between an experienced and inexperienced trainee, using either a checklist score or an overall competence rating. Vascular OR nurses achieve similar results to the surgeons, suggesting that they could also assess the operative competence of surgical trainees. This should come as no surprise, as OR nurses spend much of their time closely observing surgical procedures, and are often more familiar with them than the trainee! We only used two videos for this exercise as we felt that we could not persuade volunteers to watch for longer than 30 min. It would be interesting to repeat the study to include trainees of intermediate experience.

A previous study has demonstrated excellent inter-rater reliability between direct observation of SFD in the OR by a consultant surgeon and subsequent blinded assessment of videotape recordings using a structured checklist (Cronbach's alpha = 0.96).[7] Such assessments have also been shown to possess good criterion validity, in that more experienced trainees achieve better scores.[10] Direct observation in the OR is probably the 'gold standard' in terms of face validity as it measures 'the real thing', and it may also be better for the assessment of other competencies such as judgement, use of assistants and communication skills. However, such assessments are time-consuming as they require the presence of a consultant and they can be influenced by observer bias, which affects their reliability.[11] However, assessors can be trained to reduce the risk such bias, and assessment by nurses would avoid the need for a consultant to be present in the OR at all times.

Video recording permits subsequent fast-forwarding during uneventful parts of the operation. This can reduce the assessment time by up to 50%, without loss of much information, although this does depend upon the specific operation.[12] The good reproducibility of the video recordings of SFD is probably because it is an operation with little variation in anatomy or difficulty. Video recordings also avoid the risk of observer bias, as the recordings can be easily blinded.

Surgical simulations are useful for assessment purposes because they can be standardised.[13] However, trainees may perform differently on simulations compared to a real operation for many reasons.[14] These include unrealistic simulations, the stress caused by an examination or because of the more complex environment in the OR.[15] Therefore, it seems

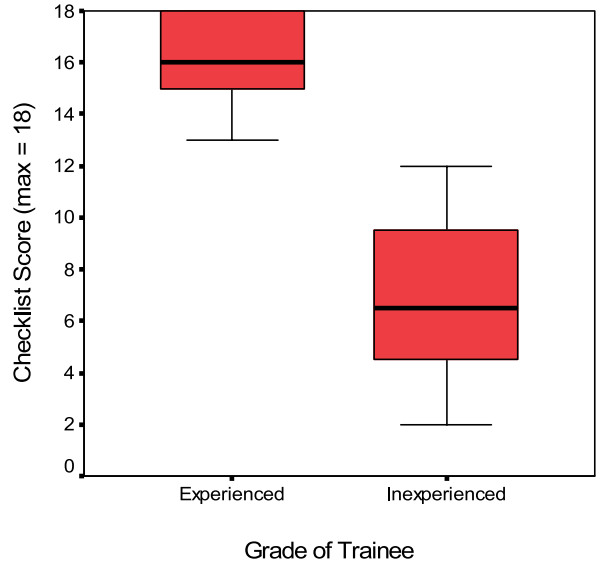

Fig. 1. Checklist scores for the experienced and inexperienced trainees by the surgical assessors (consultants and trainees). Chi-squared $p<0.0001$. Median represented by bold bar, interquartile range by box and range by whiskers.

sensible to assess trainees by both methods in order to reduce possible sampling errors. Video recordings of operations, combined with structured checklists, provide excellent feedback to trainees (formative assessment). Such recordings could also form part of a trainee's portfolio and contribute to subsequent examinations (summative assessment). Nurses could also be used to assess the operative competence of surgical trainees.

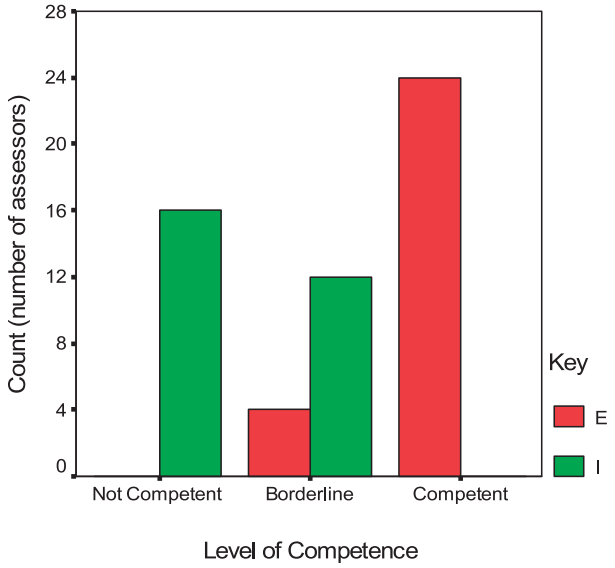

Fig. 2. Competence categories allocated the surgical assessors (consultants and trainees) for the experienced (E) and inexperienced (I) trainees. Chi-squared $p<0.0001$.

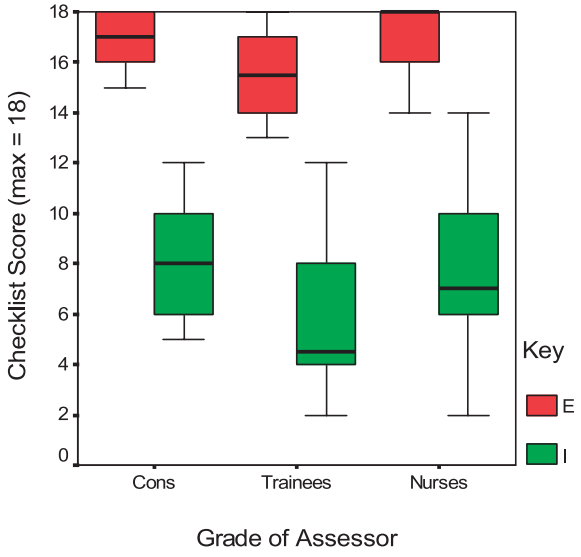

Fig. 3. Checklist scores for the experienced (E) and inexperienced (I) trainee videos by the consultants ($n=14$), trainees ($n=14$) and OR nurses ($n=13$). Chi-square $p<0.0001$ for all groups of assessors.

Acknowledgements

Our thanks go to the members of the Education and Training Committee, ordinary members of the Vascular Society and members of the Society of Vascular Nurses who participated in this study.

References

1 Galasko C. Competencies required to be a competent surgeon. *Ann R Coll Surg Engl* 2000;**82**(Suppl):89–90.
2 Department of Health. European Working Time Directive. 2003; www.doh.gov.uk/workingtime/index.htm.
3 Katory M, Singh S, Beard JD. Twenty trent trainees: a comparison of operative competence after BST. *Ann R Coll Surg Engl* 2001;**83**(Suppl):328–330.
4 Pandey VA, Wolfe JHN, Lindahl AK, Rauwerda JA, Bergqvist D, on behalf of the European Board of Vascular Surgery. Validity of an exam assessment in surgical skill EBSQ-VASC pilot study. *Eur J Vasc Endovasc Surg* 2004;**27**:341–348.
5 Anastakis DJ, Regehr G, Reznick RK, Cusimano M, Murnaghan J, Brown M *et al*. Assessment of technical skills transfer from the bench training model to the human model. *Am J Surg* 1999;**76**:167–170.
6 Martin JA, Regehr G, Reznick RK, MacRoe H, Murnaghan J, Hutchinson C. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997;**84**:273–278.
7 Beard JD, Jolly BC, Newble DI, Thomas WEG, Donelly J, Southgate LJ. Assessing the technical skills of surgical trainees. *Br J Surg* 2005; [In press].
8 Joint Committee on Postgraduate Training for General Practice. Assessment. 2004; www.jcptgp.org.uk/certification/assessment.asp.
9 Norcini JJ. Setting standards on educational tests. *Med Educ* 2003;**37**:464–469.
10 Mulkern EM, Beard JD. Saphenofemoral disconnection: do vascular trainees perform better? *Br J Surg* 2004;**91**:1077.

11 Warf BC, Donnelly MB, Schwartz RW, Sloan DA. The relative contributions of interpersonal and specific clinical skills to the perception of global clinical competence. *J Surg Res* 1999;**86**:17–23.
12 Scott DJ, Rege RV, Bergen PC, Guo WA, Laycock R, Tesfay ST *et al*. Measuring operative performance after laparoscopic skills training: edited videotape versus direct observation. *J Laparosc Adv Surg Tech* 2000;**10**:183–190.
13 Taffinder N, McManus I, Jansen J, Russell R, Darzi A. An objective assessment of surgeon's psychomotor skills: validation of the MIST-VR laparoscopic simulator. *Br J Surg* 1998; **75**(Suppl):281.
14 Paisley AM, Baldwin PJ, Paterson-Brown S. Validity of surgical simulation for the assessment of operative skill. *Br J Surg* 2001;**88**:1525–1532.
15 Rethans J-J, Norcini JJ, Baron-Maldonado M, Blackmore D, Jolly BC, LaDuca T *et al*. The relationship between competence and performance: implications for assessing practice performance. *Med Educ* 2002;**36**:901–909.

Accepted 31 January 2005
Available online 11 March 2005